

TABLE 4-continued

CRW Unigene sequences and Insect Nucleotide Sequence Homologs					
SEQ ID NO ¹	Identity Position ²	Gene ID ³	Identity Poston ⁴	% Identity ⁵	Genus species ⁶
40704	113-308	GI:14699645	95-290	85%	<i>Drosophila melanogaster</i>
40704	113-308	GI:14697887	94-289	85%	<i>Drosophila melanogaster</i>
40704	113-308	GI:14697103	136-331	85%	<i>Drosophila melanogaster</i>
40704	113-308	GI:14696099	137-332	85%	<i>Drosophila melanogaster</i>
40704	113-308	GI:14696107	136-331	85%	<i>Drosophila melanogaster</i>
40704	113-308	GI:14695238	95-290	85%	<i>Drosophila melanogaster</i>
40704	113-308	GI:14693081	133-328	85%	<i>Drosophila melanogaster</i>
40704	113-308	GI:14691490	138-333	85%	<i>Drosophila melanogaster</i>
40705	694-1364	GI:2454487	811-1481	84%	<i>Aedes aegypti</i>
40705	715-1220	GI:22039978	3-507	86%	<i>Ctenocephalides felis</i>
40705	694-1175	GI:4734043	166-647	85%	<i>Aedes aegypti</i>
40705	895-1286	GI:16899106	3-393	87%	<i>Ctenocephalides felis</i>
40705	895-1286	GI:16899780	6-395	87%	<i>Ctenocephalides felis</i>
40705	895-1286	GI:16899721	6-396	86%	<i>Ctenocephalides felis</i>
40705	961-1286	GI:22039013	8-333	87%	<i>Ctenocephalides felis</i>
40705	874-1327	GI:33376955	30-483	83%	<i>Glossina morsitans morsitans</i>
40705	636-1136	GI:46997165	360-859	81%	<i>Acyrtosiphon pisum</i>
40705	874-1220	GI:33376948	25-371	84%	<i>Glossina morsitans morsitans</i>
40705	943-1364	GI:3514814	74-495	82%	<i>Drosophila melanogaster</i>
40705	943-1364	GI:24583987	1055-1476	82%	<i>Drosophila melanogaster</i>
40705	694-884	GI:24583987	806-996	82%	<i>Drosophila melanogaster</i>
40705	943-1364	GI:24583985	967-1388	82%	<i>Drosophila melanogaster</i>
40705	694-884	GI:24583985	718-908	82%	<i>Drosophila melanogaster</i>
40705	943-1364	GI:24583983	1052-1473	82%	<i>Drosophila melanogaster</i>
40705	694-884	GI:24583983	803-993	82%	<i>Drosophila melanogaster</i>
40705	943-1364	GI:18467973	1049-1470	82%	<i>Drosophila melanogaster</i>
40705	694-884	GI:18467973	800-990	82%	<i>Drosophila melanogaster</i>
40705	943-1364	GI:19527546	1052-1473	82%	<i>Drosophila melanogaster</i>
40705	694-884	GI:19527546	803-993	82%	<i>Drosophila melanogaster</i>
40705	1045-1365	GI:4734199	1-321	84%	<i>Aedes aegypti</i>
40705	943-1280	GI:51961912	81-418	83%	<i>Drosophila simulans</i>
40705	734-947	GI:22039138	73-285	87%	<i>Ctenocephalides felis</i>
40705	959-1364	GI:24583991	1081-1486	81%	<i>Drosophila melanogaster</i>
40705	959-1364	GI:18467977	1081-1486	81%	<i>Drosophila melanogaster</i>
40705	943-1340	GI:21355198	994-1391	81%	<i>Drosophila melanogaster</i>
40705	694-884	GI:21355198	745-935	82%	<i>Drosophila melanogaster</i>
40705	959-1364	GI:19528270	1021-1426	81%	<i>Drosophila melanogaster</i>
40705	959-1364	GI:18859618	951-1356	81%	<i>Drosophila melanogaster</i>
40705	943-1340	GI:1373432	994-1391	81%	<i>Drosophila melanogaster</i>
40705	694-884	GI:1373432	745-935	82%	<i>Drosophila melanogaster</i>
40705	959-1364	GI:5851682	1021-1426	81%	<i>Drosophila melanogaster</i>
40705	142-345	GI:22039875	163-366	87%	<i>Ctenocephalides felis</i>
40705	142-345	GI:16901137	217-420	87%	<i>Ctenocephalides felis</i>
40705	82-595	GI:34787824	112-625	79%	<i>Callosobruchus maculatus</i>
40705	142-345	GI:16901267	156-360	87%	<i>Ctenocephalides felis</i>
40705	771-1022	GI:22005558	58-309	84%	<i>Aedes aegypti</i>
40705	96-357	GI:46996282	118-379	83%	<i>Acyrtosiphon pisum</i>
40705	963-1364	GI:18898890	11-412	80%	<i>Anopheles gambiae</i>
40705	967-1364	GI:18936027	25-422	80%	<i>Anopheles gambiae</i>
40705	61-344	GI:37952369	124-407	81%	<i>Ips pini</i>
40706	1230-1251	GI:33371240	247-268	100%	<i>Glossina morsitans morsitans</i>
40706	1230-1251	GI:33374947	249-270	100%	<i>Glossina morsitans morsitans</i>

¹WCR SEQ ID NO as set forth in the sequence listing;

²Nucleotide position in the SEQ ID NO in column 1 that exhibits substantial identity with Gene ID in column 3 on same row;

³Gene accession number of corresponding matching sequence identified within public database that exhibits substantial identity with column 1 SEQ ID NO;

⁴nucleotide position of sequence identified in column 3 that matches with CRW nucleotides specified on same row;

⁵Percentage identity between the WCR SEQ ID NO and Gene ID (comparison of identity between column 2 and column 4 sequences on any given row); and

⁶Genus and species of organism from which the Gene Accession No. sequence was derived.

Example 15

[0371] This example illustrates the identification of predicted protein functional domains and gene families from the translation of the nucleotide sequences disclosed herein using sequence matches to known sequences and existing domain consensus models.

[0372] The protein sequences were first produced with a “translator” program, which translated Unigenes into peptide

sequences through the following steps: homology to known proteins; model-based ab initio gene structure prediction; and longest open reading frame (ORF). Frame shifts due to sequencing errors were corrected. The protein sequences were then searched against Pfam database, a large collection of multiple sequence alignments and hidden Markov models (HMM) covering many common protein families (The Pfam Protein Families Database, Bateman et al., *Nucleic Acids*